J. Mol. Biol. (1973) 78, 363–376

Studies of Simian Virus 40 DNA

VII.† A Cleavage Map of the SV40 Genome

KATHLEEN J. DANNA‡, GEORGE H. SACK, JR AND DANIEL NATHANS

Department of Microbiology, The Johns Hopkins University School of Medicine, Baltimore, Md 21205, U.S.A.

(*Received 28 December 1972*)

A physical map of the Simian virus 40 genome has been constructed on the basis of specific cleavage of Simian virus 40 DNA by bacterial restriction endonucleases. The 11 fragments produced by enzyme from *Hemophilus influenzae* have been ordered by analysis of partial digest products and by analysis of an overlapping set of fragments produced by enzyme from *Hemophilus parainfluenzae*. In addition, the single site in SV40 DNA cleaved by the *Escherichia coli* R_I restriction endonuclease has been located. With this site as a reference point, the *H. influenzae* cleavage sites and the *H. parainfluenzae* cleavage sites have been localized on the map.

1. Introduction

The genome of the oncogenic Simian virus 40 is a double-stranded, covalently closed DNA molecule with a molecular weight of about 3×10^6. We have been studying the structure and function of this molecule by specific cleavage with bacterial restriction endonucleases (Adler & Nathans, 1970; Danna & Nathans, 1971,1972; Nathans & Danna, 1972; Sack & Nathans, 1973). Such enzymes make double strand breaks in DNA, generally at specific sites (Kelly & Smith, 1970), thus providing reference points as well as products derived from particular parts of the genome. We reported earlier that restriction endonuclease from *Hemophilus influenzae* produces eleven fragments from SV40§ DNA (Danna & Nathans, 1971) and that the enzyme from *Hemophilus parainfluenzae* produces three major fragments separable by electrophoresis (Sack & Nathans, 1973). In order to use these fragments and enyzme cleavage sites to localize genes and template functions of the SV40 genome, it is necessary to order the fragments in the molecule. By analysis of partial digest products and overlapping sets of fragments, we have ordered the SV40 DNA pieces produced by the restriction endonucleases from *H. influenzae* and *H. parainfluenzae*. In addition, the site cleaved by the *Escherichia coli* R_I restriction endonuclease (Morrow & Berg, 1972; Mulder & Delius, 1972) has also been localized. On the basis of these data and estimates of the size of each DNA fragment, we have constructed a physical map of the SV40 genome.

† Paper VI in this series is Sack & Nathans, 1973.

‡ Present address: Laboratorium voor Moleculaire Biologie, K.L. Ledeganckstraat 35, Rijksuniversiteit-Gent, Gent 9000, Belgium.

§ Abbreviation used: SV40, simian virus.

2. Nomenclature for DNA fragments

In this paper we have used a simplified nomenclature to designate fragments of SV40 DNA produced by restriction endonucleases. Each fragment present in a complete digest is assigned a capital letter in order of fragment size (A, B, C . . . where A is largest) with an italicized prefix designating the enzyme used to produce the fragment. The restriction endonuclease of *H. influenzae* (Endo R of Smith & Wilcox, 1970) is denoted as *Hin*, that from *H. parainfluenzae* (Gromkova & Goodgal, 1972) as *Hpa*, and the *E. coli* R_I enzyme (Yoshimoro, 1971) as Eco_{RI}. Thus the SV40 DNA fragments found in an *H. influenzae* enzyme digest are *Hin*-A, *Hin*-B, etc., and those found in an *H. parainfluenzae* enzyme digest are *Hpa*-A, *Hpa*-B, etc. When a new fragment results from sequential digestion by two enzymes, it is denoted by combining the designations of the single digest fragments from which it originates. For example, as will be shown below, sequential digestion of SV40 DNA with restriction endonucleases of *H. parainfluenzae* and *H. influenzae* yields a new fragment, not present in either single digest, which represents the overlap between fragments *Hpa*-B and *Hin*-C. The new fragment is designated *Hpa*-B *Hin*-C or *Hin*-C *Hpa*-B (Fig. 1(a)).

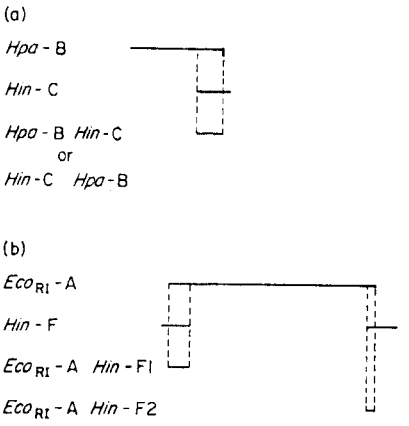

FIG. 1. Examples of nomenclature for new fragments resulting from sequential digestion of SV40 DNA by 2 restriction endonucleases. Horizontal lines indicate fragments and vertical lines indicate cleavage sites.

In those cases in which one of the restriction enzymes cleaves a circular molecule once, or, more generally, only between two adjacent cleavage sites of a second enzyme used subsequently, the above nomenclature is ambiguous; sequential cleavage by the second enzyme yields a new fragment from each end of the first product. We have labeled these fragments 1 and 2, fragment 1 being the larger. For example, the R_I restriction endonuclease from *E. coli* cleaves SV40 DNA at a single unique site within *Hin*-F (see below). Subsequent cleavage of this full-length linear product (Eco_{RI}-A) by the *H. influenzae* enzyme would be expected to yield two new products consisting of the left and right parts of *Hin*-F. These are designated Eco_{RI}-A *Hin*-F1 for the larger product and Eco_{RI}-A *Hin*-F2 for the smaller product (Fig. 1(b)).

In some instances, an isolated fragment produced by one enzyme may be further cleaved by a second enzyme into two or more products. Until the origin of each final

product is known it is not possible to designate these fragments by the above nomenclature. In this instance, we suggest a less specific designation indicating the origin of the isolated fragment and the second enzyme used. For example, the *H. parainfluenzae* enzyme cleaves the isolated *Hin*-C fragment into two products. Before identifying the *Hpa* fragments from which the two products are derived, we designate them *Hpa Hin*-C1 and *Hpa Hin*-C2, respectively, *Hpa Hin*-C1 being the larger.

In addition to this nomenclature for DNA fragments, we have used arabic numerals to designate restriction enzyme cleavage *sites* in the DNA, again with the prefix indicating the enzyme used. For example, *H. influenzae* cleavage sites are *Hin*-1, *Hin*-2, etc.

As many more restriction endonucleases from a single species come into use, it will be necessary to add a qualifying term to the enzyme designation, as is already the case for *E. coli* enzymes. We suggest a subscript be used, as done in this paper for the *E. coli* R_I enzyme (Eco_{RI}).

3. Materials and Methods

(a) *Cell lines and virus*

Small-plaque SV40 (isolated from strain 776 by K. Takemoto) was grown from a cloned stock in the BSC-1 line of African green monkey kidney cells, as described previously (Danna & Nathans, 1971). Stocks were made by infecting at a multiplicity of 0·001 plaque-forming units/ml.

(b) *Preparation of ^{32}P-labeled SV40 covalently closed circular DNA*

Procedures for infection of cells with SV40, and for labeling and purification of covalently closed SV40 DNA, have been given elsewhere (Danna & Nathans, 1971). Briefly, BSC-1 cells were infected at a multiplicity of 10 to 20 plaque-forming units/ml, labeled with [^{32}P]orthophosphate, and lysed by the method of Hirt (1967). Virus DNA was purified from the supernatant fluid by incubation with heated ribonuclease A, phenol extraction, ethanol precipitation, equilibrium centrifugation in CsCl/ethidium bromide, and sedimentation in a neutral sucrose gradient. Final spec. act. of ^{32}P-labeled covalently closed circular SV40 DNA was 2×10^5 to 3×10^5 cts/min/μg.

(c) *Cleavage of SV40 DNA with* H. influenzae *restriction endonuclease*

Two enzyme preparations purified by the method of Smith & Wilcox (1970) were used for most of the experiments to be described. A third preparation of enzyme used in some of the later experiments was supplied by H. O. Smith. Enzyme units have been defined by Smith & Wilcox (1970). For preparation of complete digests, SV40 covalently closed circular DNA was incubated with *H. influenzae* restriction endonuclease at 37°C in a reaction mixture containing 6·6 mM-Tris (pH 7·5), 6·6 mM-$MgCl_2$, 50 mM-NaCl. With each new preparation of enzyme, a series of enzyme concentrations was used to determine the optimum conditions for complete digestion of SV40 DNA. Some variation was found with different preparations and often high enzyme concentrations were inhibitory. Generally, incubation of 50 μg of covalently closed circular SV40 DNA with 0·066 unit of restriction endonuclease in a volume of 0·12 ml for 1 h resulted in complete digestion, as determined by electrophoresis of the products (see below). All reactions were terminated by addition of EDTA to 0·02 M.

To obtain incomplete digests of ^{32}P-labeled SV40 DNA, both the ratio of enzyme to DNA and the time of incubation were varied, but in all cases incubation was carried out at 37°C in the standard buffer described above. As an example, 0·004 unit of enzyme incubated with 2·3 μg of DNA in a volume of 0·12 ml for 30 min yielded at least 29 intermediates in addition to some limit products of the reaction. The legend to Plate I describes other conditions used for partial digestion. Individual, partially digested fragments (at concentrations ranging from 0·5 to 300 μg/ml) were completely digested by incubation with excess enzyme under standard conditions in volumes of 20 to 30 μl.

(d) *Cleavage of SV40 DNA with* H. parainfluenzae *restriction endonuclease*

The enzyme preparation used for these experiments was purified as described earlier (Gromkova & Goodgal, 1972; Sack & Nathans, 1973) and stored at 4°C in 25% glycerol. It contained 8·5 units/ml. Incubations were carried out for 90 min at 30°C in buffer containing 13 mM-Tris·HCl (pH 7·5), 20 mM-KCl, 5 mM-$MgCl_2$, 13 mM-mercaptoethanol, 3% bovine serum albumin and 0·024 unit of enzyme/μg of DNA. Under these conditions the DNA was completely digested, as judged by electrophoresis and radioautography.

(e) *Cleavage of SV40 DNA with* E. coli R_I *restriction endonuclease*

The enzyme preparation used was generously provided by H. Boyer. It was stored at −15°C in 30% glycerol. Incubations were carried out for 2 to 16 h at 30°C in 0·1 M-Tris·HCl (pH 7·6), 6 mM-$gMCl_2$, 6 mM-mercaptoethanol, 200 μg bovine serum albumin/ml and 20 μg of enzyme/μg of DNA. Conversion of covalently closed circular DNA to full length linear molecules was confirmed by centrifugation of the product together with marker SV40 DNA in 5% to 20% sucrose gradients containing 1 × SSC (SSC is 0·15 M-NaCl, 0·015 M-sodium citrate), 10 mM-Tris·HCl (pH 7·4) at 49,000 revs/min in a SW50 Spinco rotor at 20°C for 3 h.

(f) *Vertical slab gel electrophoresis*

Before electrophoresis, samples of digested DNA were made 1% in sodium dodecyl sulfate, incubated at 37°C for 30 min and 0·25 vol. of 75% sucrose, 1% bromphenol blue was added. Electrophoresis was carried out at a constant voltage of 160 V for 16 to 20 h in 4% polyacrylamide vertical slab gels (15 cm × 40 cm × 0·16 cm) in a buffer containing 40 mM-Tris, 20 mM-sodium acetate, 2 mM-sodium EDTA, pH 7·8. A gel chamber similar to those described by DeWachter & Fiers (1971) and by Reid & Bieleski (1968) was used. Preparative electrophoresis of digestion products was followed by radioautography of the wet gel to locate [^{32}P]DNA bands. Gel segments corresponding to these bands were excised at room temperature and eluted with 2 successive vol. of 0·4 ml of 0·1 × SSC, pH 7·4. In preparation for redigestion, eluates were concentrated to dryness by evaporation, dissolved in 20 to 40 μl of sterile deionized water, and dialyzed against 0·1 × SSC, pH 7·4. For radioautographic analysis of DNA fragments, gels were dried by Maizel's modification (1971) of the method described by Fairbanks *et al.* (1965) and then placed in direct contact with Kodak blue medical X-ray film. Satisfactory radioautograms were obtained with as little as 100 cts/min of a ^{32}P-labeled fragment after exposure for 5 days, the band area being about 6 mm^2.

4. Results

(a) *Order of SV40 DNA fragments produced by* H. influenzae *restriction endonuclease*

We have reported earlier that cleavage of SV40 DNA by *H. influenzae* restriction endonuclease yields 11 fragments (designated *Hin*-A to *Hin*-K) separable by polyacrylamide gel electrophoresis (Plate 1(a)). On the basis of the relative yield of each fragment from uniformly labeled [^{32}P]DNA, their molecular weights have been estimated (Table 1), and for fragments A to F, these values have been confirmed by electron micrographic length measurements (Danna & Nathans, 1971). Although it has not been established directly that the smaller fragments (G to K) are present in unimolar amount, we shall assume that this is the case; as shown later, none of the fragments is present in more than one location in the SV40 DNA molecule.

The general approach for ordering DNA fragments to be presented in this section was the separation by polyacrylamide gel electrophoresis and localization by radioautography of individual ^{32}P-labeled fragments incompletely digested with *H. influenzae* restriction enzyme; subsequently, each partial product was eluted and

TABLE 1

Molecular weight estimates of SV40 DNA fragments produced by cleavage with H. influenzae *restriction endonuclease*

Hin fragment	Relative molecular weight (% of SV40 DNA)
A	22·5
B	15·0
C	10·5
D	10·0
E	8·5
F	7·5
G	7·0
H	5·5
I	5·0
J	4·5
K	4·0

Estimates are based on the distribution of radiolabel in a complete *Hin* digest of ^{32}P-labeled SV40 DNA. For A to F, the values were confirmed by electron micrographic measurements (Danna & Nathans, 1971).

digested with excess enzyme and the ultimate products identified by their electrophoretic mobilities. In this way, several overlapping fragments were analyzed and the order of nearly all 11 of the *Hin* fragments was deduced.

Partial digestion products of ^{32}P-labeled SV40 DNA were prepared by incubating the DNA with *H. influenzae* enzyme under standard conditions for short periods or with small quantities of enzyme for longer periods. Either procedure led to the accumulation of numerous intermediate products, identified by electrophoresis and radioautography, the average size of which depended on the time of incubation and enzyme concentration. Typical results, illustrating the effect of incubation time, are shown in the radioautogram in Plate I(b). As seen in the Plate (samples 1, 2 and 3) several discrete fragments are present in addition to the limit digestion products. Also, it is evident from these experiments and numerous others that some of the limit products (fragments A, E and K) appeared quite early during the course of the reaction, indicating that certain of the susceptible sites in SV40 DNA are cleaved preferentially under these conditions. (With some preparations of enzyme, a different set of fragments (G and J) consistently appeared early, suggesting that the *H. influenzae* endonuclease preparation contained enzymes of more than one specificity (see Discussion).)

Several large-scale partial digests of ^{32}P-labeled SV40 DNA were prepared under different incubation conditions and the products purified and redigested as described in Materials and Methods and in the legend to Plate I. For radioautographic analysis, each redigested partial product was electrophoresed in a slab gel in parallel with a sample of the original undigested partial product and a complete digest marker. The R_F for each undigested partial product was determined relative to fragment G, and the limit products derived after redigestion were identified by visual comparison with the marker. Several examples of these analyses are presented in Plate I ((c), (d) and (e)). In the examples of Plate I(c), sample 2 yielded fragments B, F, G, J and K upon redigestion (sample 2d); samples 3 and 4 each yielded fragments B, F, G and

J after redigestion (samples 3d and 4d); sample 6 produced fragments C and D upon redigestion (sample 6d); and sample 7 produced fragments H and I (7d). The sample shown in Plate I(d) yielded fragments B, G and J; sample 1 of Plate I(e) yielded fragments E and K; and sample 2 of Plate I(e) produced fragments G and J after redigestion. In some instances, the putative partial products isolated from preparative gels were actually limit products. For example, in Plate I(c) sample 5 from the partial digest had the same mobility as the marker fragment A, and after redigestion, yielded no other fragments. The same result was obtained with putative partial products that had mobilities identical to those of fragments B, C, D, E, G, H and K, confirming that these fragments do not contain additional cleavage sites for the *H. influenzae* restriction endonuclease.

Some undigested partial products were clearly a mixture of two fragments of approximately equal length, as shown by electrophoresis of the limit products, and in some cases by re-electrophoresis of the partial product. Upon redigestion, two groups of limit products were formed, readily distinguished by the intensity of bands in the radioautogram. In the example shown in Plate I(c), sample 1, fragments A, C and D constitute one group; fragments B, F, G, H, I and J form the other.

As a further check on the composition of the partial products, we compared the molecular weight of each partially digested fragment with the sum of the molecular weights of the products derived from it. For these calculations we used the molecular weight values for the limit products shown in Table 1. Molecular weights of the partial products were estimated graphically from a plot relating log molecular weight

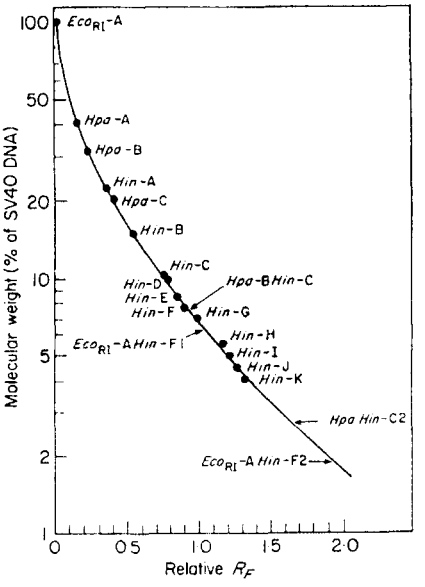

FIG. 2. Log molecular weight *versus* electrophoretic mobility of DNA fragments in 4% polyacrylamide gel (5% cross-linking). Mol. wt is expressed as a percentage of the mol. wt of SV40 DNA, and mobility is given relative to fragment *Hin*-G

$$\left(\frac{R_F \text{ of fragment}}{R_F \text{ of } Hin\text{-G}}\right),$$

For each fragment, the mol. wt has been estimated by the yield of each fragment in a total *Hin* or *Hpa* digest. The arrows indicate the mobility of new fragments resulting from sequential cleavage by 2 restriction endonucleases (see text).

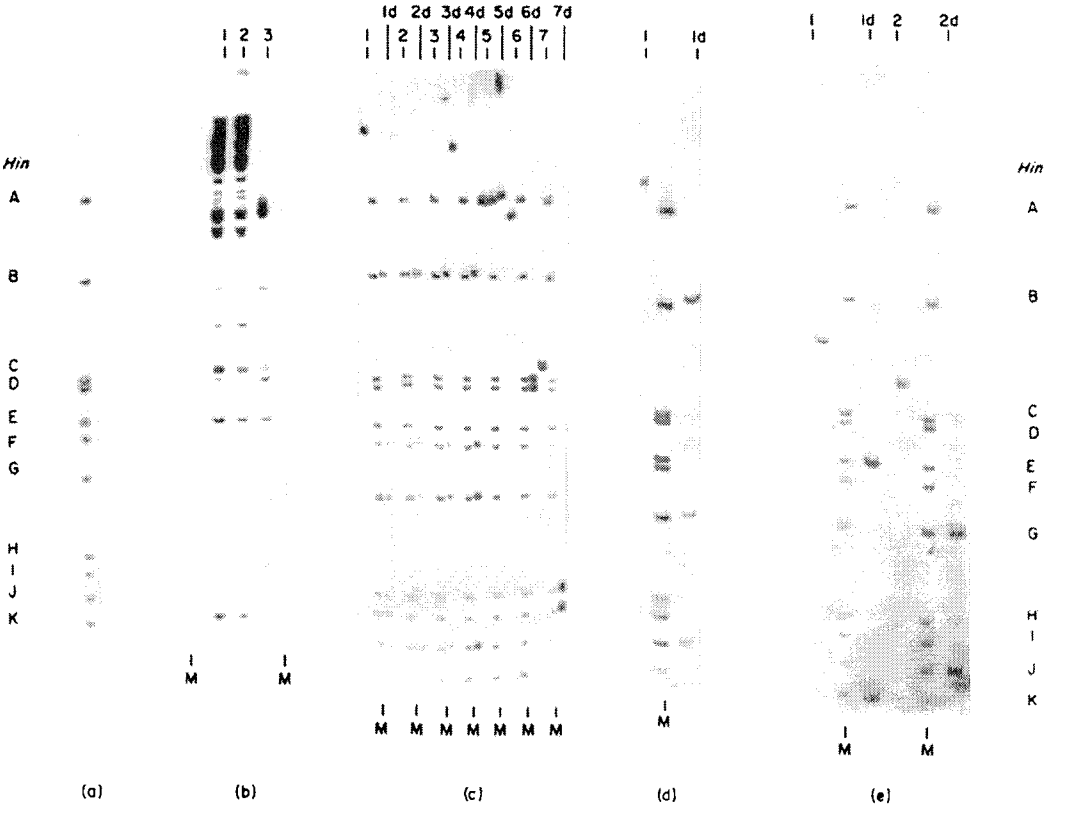

PLATE I. Examples of partial and completed digests of ^{32}P-labeled SV40 DNA with restriction endonuclease from *H. influenzae*. Each plate is a radioautogram of a single gel slab following electrophoresis. The origin is at the top.

(a) A complete digest of SV40 DNA. Conditions for digestion, electrophoresis and radioautography are described in Materials and Methods. Electrophoresis was for 17·5 h at 150 V.

(b) Partial digests of SV40 DNA. 25 μg of ^{32}P-labeled SV40 DNA I (10^5 cts/min/μg) were incubated in a vol. of 0·1 ml with 0·016 of a unit of enzyme under standard conditions. Samples were removed after 20 min digestion (sample 2) and after 30 min digestion (sample 1). For sample 3, 23 μg of ^{32}P-labeled SV40 DNA I ($4{\cdot}5 \times 10^4$ cts/min/μg) were incubated for 1 h with 0·055 of a unit of enzyme in a vol. of 0·32 ml. Samples of the partial digests (1, 2 and 3) and of a complete digest marker (M) were electrophoresed for 16 h at 160 V.

(c), (d) and (e) Redigestion of partial digest products eluted from electrophoresis gels. For each sample, the partial product (1, 2 etc.) and the redigested partial product (1d, 2d etc.) were electrophoresed in the sample gel slab with a complete digest of SV40 DNA as marker (M). (c) Electrophoresis was carried out at 165 V for 17 h; (d) 130 V for 21 h and 180 V for an additional 3 h; (e) 150 V for 17·5 h.

[*facing p. 368*

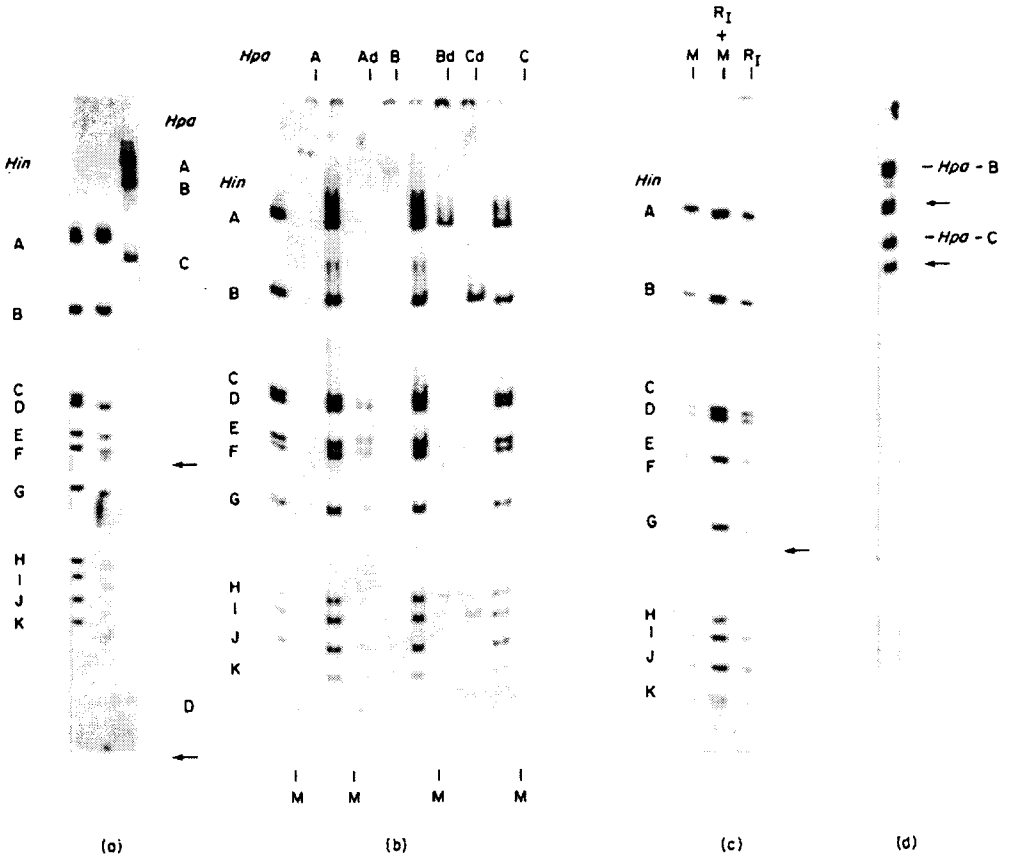

PLATE II. Analysis of ^{32}P-labeled SV40 DNA digests produced by restriction endonuclease from *H. parainfluenzae* and from *E. coli* R_I.

(a) Digestion of SV40 DNA with *H. influenzae* and *H. parainfluenzae* enzymes. The column on the left is a *Hin* digest, that on the right is a *Hpa* digest, and the middle column is a digest produced by sequential incubation with *Hpa* and *Hin* enzymes. Conditions of electrophoresis are given in the legend to Plate I(a). (The slightly faster mobility of some of the double digest products is probably related to the higher salt concentration in this sample—see Plate II(b).) The arrows indicate fragments *Hpa Hin*-C1 (upper) and *Hpa Hin*-C2 (lower).

(b) Digestion of isolated *Hpa* fragments with *H. influenzae* enzyme. The undigested *Hpa* fragments (A, B and C) and the digested *Hpa* fragments (Ad, Bd and Cd) were electrophoresed in a 4% slab gel with complete *Hin* digest marker (M) for 20 h at 145 V.

(c) Digestion of Eco_{RI}-A with *H. influenzae* enzyme. Samples were electrophoresed for 18 h at 155 V. The column on the left (M) is a digest of covalently closed circular SV40 DNA with *H. influenzae* enzyme; the column on the right (R_I) is a digest of Eco_{RI}-A with *H. influenzae* enzyme; and the middle column (R_I + M) is a mixture of the 2 digests. The arrow designates a new fragment seen only in the R_I sample. (The expected second new fragment was not observed; see text.) Conditions of electrophoresis as in (b).

(d) Digestion of Eco_{RI}-A with *H. parainfluenzae* enzyme. The location of the *Hpa* fragments in another part of the gel is indicated. Conditions of electrophoresis as in (b). The arrows indicate the 2 new fragments Eco_{RI}-A *Hpa*-A1 (top) and Eco_{RI}-A *Hpa*-A2 (bottom).

to R_F in 4% polyacrylamide gels, using as standards of known molecular weight the *Hin* limit products A to D (Danna & Nathans, 1971); unit-length linear SV40 DNA produced by cleavage with the *E. coli* R_I restriction endonuclease (Morrow & Berg, 1972; Mulder & Delius, 1972); and two fragments produced by the restriction endonuclease from *H. parainfluenzae*, fragments *Hpa*-A (relative mol. wt, 42%) and *Hpa*-B (relative mol. wt, 34%) (Sack & Nathans, 1973) (Fig. 2). Although the curve increases markedly in slope above a relative molecular weight of about 30% of the SV40 DNA, it was nonetheless possible to estimate the sizes of large partial products in this way.

Table 2 summarizes the data from all partial products analyzed. The results were generally straightforward, the molecular weight of each partial digestion product being identical to or close to the sum of the molecular weights of the products derived from it. Slight discrepancies might be due to the non-linearity of the relationship between log molecular weight and R_F, particularly for larger fragments, and the consequent difficulty in estimating molecular weights.

To deduce the physical order of the *Hin* products from the data in Table 2, we have arranged the limit products derived from each incompletely digested fragment

TABLE 2

Redigestion of partial digestion products with H. influenzae *restriction endonuclease*

	Undigested partial product		Redigested partial product	
R_F	Estimated molecular weight (% of SV40 DNA)	Products	Sum of product molecular weights (% of SV40 DNA)	Overlapping orders
0·69	12·0†	G,J	12·0	G J
0·68	11·5†	H,I	11·5	(H,I)
0·66	12·0†	F,K	11·5	F K
0·62	13·0†	E,K	12·5	K E
0·59	13·5†	F,J	12·5	J F
0·39	20·3†	F G,J	19·5	G J F
0·34	22·0†	C,D	20·5	(C,D)
	22·0†	B,G	22·0	B G
0·27	24·5†	B,G,J	27·0	B G J
0·19	36·0‡	B,F,G,J	34·5	B G J F
0·17	40·0‡	B,F,G,J,K	38·5	B G J F K
0·17	40·0‡	A,C,D	43·0	(C,D) A
0·16	43·0‡	A,C,D	43·0	(C,D) A
	43·0‡	B,F,G,H,I,J	45·0	(H,I) B G J F
0·14	48·0‡	B,F,G,H,I,J,K	49·0	(H,I) B G J F K
0·13	51·0‡	A,C,D,E	51·5	E (C,D) A
0·11	57·0‡	A,B,G,H,I,J	60·0	A (H,I) B G J

The R_F for each partial product was determined relative to fragment G in the complete digest in the same gel and is expressed as

$$\frac{\text{distance of partial product from origin}}{\text{distance of G from origin}}$$

When 2 readily distinguishable groups of final products were derived from a partial product, the groups are shown on 2 separate lines in the Table (R_F 0·16 and R_F 0·34).

† The *Hin* products A to D were the standards of known mol. wt in the plot relating log mol. wt to R_F used to estimate these values.

‡ The *Hpa* products A and B and full length linear SV40 DNA were the standards of known mol. wt in the plot relating log mol. wt to R_F used to estimate these values.

in overlapping positions, as shown in the last column of the Table. In parentheses are two groups, namely the pairs (C,D) and (H,I) whose positions could not be uniquely determined from these data. The order of the *Hin* products established from analysis of partial digestion products is therefore E (C,D) A (H,I) B G J F K, with E and K being contiguous in the circular molecule.

(b) *Order of SV40 DNA fragments produced by* H. parainfluenzae *restriction endonuclease*

We have reported recently (Sack & Nathans, 1973) that a restriction endonuclease from *H. parainfluenzae* (Gromkova & Goodgal, 1972) cleaves SV40 DNA into three predominant fragments: *Hpa*-A (about 42% of the length of SV40 DNA), *Hpa*-B (34%) and *Hpa*-C (20%) (Plate II(a)). To orient these large fragments in the SV40 DNA molecule, we have determined which *Hin* fragments are present in each of the *Hpa* fragments. At the same time, the results of this analysis have resolved the ambiguities in the order of *Hin* fragments.

When ^{32}P-labeled SV40 DNA I was digested successively with restriction enzymes from *H. parainfluenzae* and from *H. influenzae* and the double digestion products separated by electrophoresis, fragments corresponding to all of the *Hin* fragments except *Hin*-C were observed (Plate II(a)). In addition, two new fragments appeared, one between *Hin*-F and *Hin*-G with a mobility of 0·80 relative to *Hin*-G and the other with a mobility of 1·65 relative to *Hin*-G. It appeared, therefore, that one of the *Hpa* sites was within fragment *Hin*-C, and the two new fragments were derived by cleavage of *Hin*-C. Direct evidence for this conclusion was obtained by digesting electrophoretically purified *Hin*-C with the *H. parainfluenzae* enzyme, which resulted in the appearance of both of the new fragments just described. On the basis of electrophoretic mobilities (Fig. 2) we estimate the size of the larger product (*Hpa Hin*-C1) as 8·0% and the smaller product (*Hpa Hin*-C2) as 2·5% of the SV40 DNA molecule.

In order to determine which *Hin* fragments were contained in particular *Hpa* fragments, we isolated fragments *Hpa*-A, *Hpa*-B and *Hpa*-C following electrophoresis in 3% acrylamide gel and digested them individually with *H. influenzae* enzyme. Fragment *Hpa*-A yielded *Hin*-D, E, F, G, J, K; fragment *Hpa*-B yielded *Hin*-A and H and the new fragment *Hpa Hin*-C1; and fragment *Hpa*-C yielded *Hin*-B and I (Plate II(b) and Table 3). Therefore, we can conclude that one of the *Hpa* sites is within fragment *Hin*-C and the other two sites are identical to, or very close to, the *Hin* sites between *Hin*-H and I and between *Hin*-B and G (Table 3). (As noted in Table 3, although we expected to detect the small fragment derived from *Hin*-C (*Hpa Hin*-C2) in the *H. influenzae* enzyme digest of *Hpa*-A, we did not observe this product. This could be due to the presence of a fourth *Hpa* site near the *Hin*-C–*Hin*-D junction: see Discussion.)

By combining the data in Table 3 with those in Table 2, the orientation of the large *Hpa* fragments in the SV40 DNA molecule can be deduced and the uncertainties in the order of *Hin* fragments resolved. Ignoring the smaller part of fragment *Hin*-C that is unaccounted for, we can relate the two sets of fragments as follows:

Hin E D C A H I B G J F K
| | |
Hpa A B C A

TABLE 3

Digestion of H. parainfluenzae *restriction endonuclease products with* H. influenzae *restriction endonuclease*

Product	R_F	Estimated molecular weight (% of SV40 DNA)	Products	Sum of product molecular weights (% of SV40 DNA)	Order
Hpa-A	0·17	42	D,E,F,G,J,K	41·5	G J F K E D
Hpa-B	0·21	34	A,H,C1	34	C A H
Hpa-C	0·39	20	B,I	20	I B
					C A H I B G J F K E D

The R_F was determined relative to fragment *Hin*-G in the complete *Hin* digest in the same gel and is expressed as

$$\frac{\text{distance of } Hpa \text{ fragment from origin}}{\text{distance of } Hin\text{-G from origin}}.$$

Molecular weights of *Hpa* fragments were estimated by relative yield (Sack & Nathans, 1973). Although we expected to find the product *Hpa Hin*-C2 in the digest of *Hpa*-A, this small fragment was not observed.

(c) *The site of cleavage of SV40 DNA by* E. coli R_I *restriction endonuclease*

The *E. coli* R_I restriction endonuclease (Yoshimoro, 1971) has been shown to produce one specific double-strand break in SV40 DNA (Morrow & Berg, 1972; Mulder & Delius, 1972; Fareed *et al.*, 1972) which is valuable as a reference point in the circular molecule. For this reason, we thought it worthwhile to localize this site on the cleavage map. We first determined which *Hin* fragment contains the Eco_{RI} site by comparing an *H. influenzae* enzyme digest of covalently closed circular SV40 DNA with a digest of Eco_{RI}-A (i.e. the full length linear product obtained by cleavage of covalently closed circular DNA with the R_I enzyme). Electrophoretograms of the two digests are shown in Plate II(c). The only difference in the two digests was the absence of *Hin*-F among the Eco_{RI}-A products and the appearance of a new fragment, between *Hin*-G and *Hin*-H. It is evident, therefore, that the R_I site is within *Hin*-F. This has been confirmed by direct cleavage of *Hin-F* with the Eco_{RI} enzyme, the two products (Eco_{RI}-A *Hin*-F1 and F2) being identified electrophoretically. From their electrophoretic mobilities of 1·05 and 1·93, relative to *Hin*-G (Fig. 1), we can estimate the size of Eco_{RI}-A *Hin*-F2 as 6·0 to 6·5% and Eco_{RI}-A *Hin*-F2 as 1·5 to 2·0% of the SV40 DNA molecule.

To determine which end of fragment *Hin*-F is closest to the R_I site, we needed to isolate larger pieces of DNA from Eco_{RI}-A and determine whether fragment Eco_{RI}-A *Hin*-F1 is contiguous to fragment J or fragment K, the two *Hin* fragments adjacent to *Hin*-F. This was done by cleaving Eco_{RI}-A with *H. parainfluenzae* enzyme (Plate II(d)), isolating the two new pieces (derived from *Hpa*-A and designated Eco_{RI}-A *Hpa*-A1 and A2), and cleaving each of the latter pieces with *H. influenzae* enzyme. As shown in Table 4, fragment Eco_{RI}-A *Hpa*-A1 yielded *Hin*-K, E and D, whereas fragment Eco_{RI}-A *Hpa*-A2 yielded *Hin*-G, J, and F1. Therefore, fragment Eco_{RI}-A *Hin*-F1 is next to *Hin*-J and the SV40 cleavage site of the R_I enzyme is 6·0 to 6·5% of the length of SV40 DNA from the F–J junction within fragment *Hin*-F.

It should be noted that in the experiments involving cleavage of Eco_{RI} products by *H. influenzae* enzyme, the expected smaller *Hin*-F2 fragment was not observed, although cleavage of *Hin*-F by the Eco_{RI} enzyme did yield both pieces of *Hin*-F. At present we have no explanation for the failure to observe the *Hin*-F2 fragment in *H. influenzae* digests of Eco_{RI}-A *Hpa*-A1.

TABLE 4

Analysis of fragments Eco_{RI}-A Hpa-*A1 and* Eco_{RI}-*A* Hpa-*A2 with* H. influenzae *restriction endonuclease*

Eco_{RI}-A *Hpa* product	R_F	Estimated† molecular weight (% of SV40 DNA)	*Hin* products	Sum of product molecular weights (% of SV40 DNA)
Eco_{RI}-A *Hpa*-A1	0·29	26·0	K E D	(22·5 to 26·5)‡
Eco_{RI}-A *Hpa*-A2	0·46	17·0	G J F1	17·5

The R_F for each product was determined relative to fragment *Hin*-G in the same gel.

† The mol. wts were estimated graphically from a plot relating log mol. wt and R_F (Fig. 1).

‡ In addition to *Hin*-K, E and D (mol. wt 22·5%) Eco_{RI}-A *Hpa*-A1 should have a small part of *Hin*-F (equal to 1·5% of SV40 DNA) and possibly a small part of *Hin*-C (equal to 2·5% of SV40 DNA), neither of which was detected in this experiment.

(d) *A cleavage map of the SV40 genome*

We can usefully summarize all the foregoing data in a single cleavage map of the SV40 genome, incorporating the various cleavage sites and molecular weight estimates of the fragments (Fig. 3). Since the *E. coli* R_I enzyme site appears to be unique, this site has been designated the zero point, and measurements have been made arbitrarily in the direction F-J-G-B- ... from this point. Map distances are shown in Figure 3 as fractions of the length of SV40 DNA, and the position of each cleavage site is given in these units in Table 5.

TABLE 5

Map distances for Hin *and* Hpa *cleavage sites*

Cleavage site	Distance from Eco_{RI} site (map units or fractional length of SV40 DNA)
Eco_{RI}	0
Hin 1	0·060
Hin 2	0·105
Hin 3	0·175
Hin 4	0·325
Hin 5	0·375
Hin 6	0·430
Hin 7	0·655
Hin 8	0·760
Hin 9	0·860
Hin 10	0·945
Hin 11	0·985
Hpa 1	0·175
Hpa 2	0·375
Hpa 3	0·735
(*Hpa* 4	0·760)

5. Discussion

A cleavage map of a DNA molecule, such as that shown in Figure 3, is a physical map based on specific sites susceptible to endonucleases. The particular map of the SV40 genome shown in the Figure is based on the sites of cleavage of SV40 DNA by three bacterial restriction endonucleases: the enzymes from *H. influenzae*, *H. parainfluenzae*, and *E. coli* R_I. As other specific restriction endonucleases are used on SV40 DNA, it will be possible to map new fragments in relation to these sites.

The methods we have used to order DNA fragments are analogous to those used for ordering peptides in a protein molecule or in sequencing RNA, i.e. analysis of overlapping fragments obtained either by partial digestion with a specific cleaving enzyme or by digestion with a second enzyme of different specificity. Although in the case of proteins and RNA, amino acid or nucleotide sequence is generally used to identify overlapping oligomers, we were able to identify overlapping SV40 DNA fragments by the distinctive electrophoretic mobility of their final digestion products. In the case of more complex digests, it will be necessary to cross-hybridize DNA fragments or to determine partial nucleotide sequences in order to localize the fragments in the cleavage map. Alternatively, DNA fragments can be ordered by electron microscopic examination of heteroduplexes formed between specific fragments or between a

fragment and the full-length linear molecule. The ability to visualize the duplex region, however, sets a lower limit to the length of a fragment which can be ordered in this way. In the case of SV40 DNA, specific heteroduplexes between *H. influenzae* fragments and the linear Eco_{RI} product have been observed (T. J. Kelly, Jr, personal communication). Hybridization of denatured *Hin*-F and Eco_{RI} product yielded single-stranded circular molecules, thus confirming the localization of the Eco_{RI} site within *Hin* F.

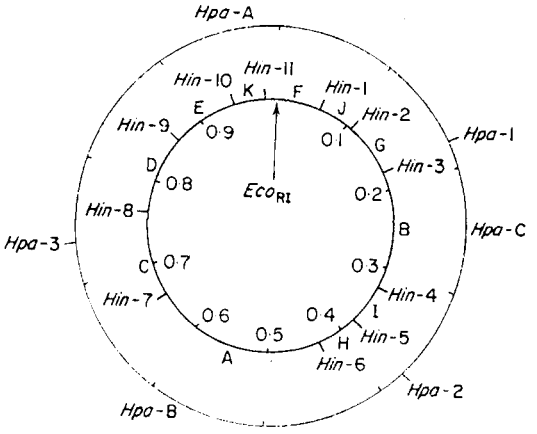

FIG. 3. A cleavage map of the SV40 genome. Map units are given as

$$\frac{\text{distance from } Eco_{RI} \text{ site}}{\text{length of SV40 DNA}}$$

in the direction F-J-G-B . . ., as given in Table 5. A fourth *Hpa* site (*Hpa*-4) probably corresponds to *Hin*-8 (see text).

Several features of the SV40 cleavage map presented in Figure 3 and Table 5 deserve comment. First, the Eco_{RI} site has been selected as the zero point of the map, since the unique full-length linear SV40 DNA molecule produced by the R_I enzyme is being used in mapping by electron microscopy. This site has been localized by identifying the *Hin* fragment in which the cleavage occurs (*Hin*-F), determining the nearest neighbor of Eco_{RI}-A *Hin*-F1 (the larger double digest product), and then estimating the molecular weight of this fragment by electrophoretic mobility.

Second, the *Hin* fragments, which have all been localized, are each present at only one position in the molecule. Earlier, we had suggested that fragments *Hin*-I, J and K appeared to be multiple on the basis of their relative electrophoretic mobility and yield (Danna & Nathans, 1971). However, with better separation of the fragments on long gel slabs, a more accurate yield of each small fragment could be determined than previously reported, and no sharp discontinuities in the yield *versus* mobility plot were found (Fig. 2). Moreover, in the absence of independent estimates of molecular weights, it is not clear that the relationship between log molecular weight and electrophoretic mobility is uniform for small pieces of DNA. In view of the unique map position for each *Hin* fragment, we now assume that the smaller fragments are actually equimolar with the other fragments. Definite proof for this assumption will require independent determinations of molecular weights, which are now being carried out.

A third feature of the map we should like to note is the coincidence of some of the *Hpa* sites with *Hin* sites. Although we cannot now exclude the possibility that these apparently coincident sites are merely very close, it seems more likely that the nucleotide sequences recognized by the *H. parainfluenzae* enzyme may include a subset of those sequences recognized by the *H. influenzae* enzyme(s) (Kelly & Smith, 1970). As pointed out earlier, in addition to the three *Hpa* sites shown in Figure 3, there may be a fourth *Hpa* site near *Hin*-8, which could account for the failure to detect fragment *Hpa Hin*-C2 from *Hpa*-A and also account for the small fragment (*Hpa*-D) present in the *H. parainfluenzae* digest of SV40 DNA (Sack & Nathans, 1973). This possibility is being investigated with more highly purified *H. parainfluenzae* endonuclease.

We should also like to comment on the finding that SV40 DNA has certain sites that are preferentially cleaved by the *H. influenzae* enzyme. Recently, the phosphocellulose fraction of *H. influenzae* restriction endonuclease has been separated into two subfractions which appear to have different specificities (H. O. Smith, personal communication). One of these fractions cleaves SV40 DNA into six fragments including *Hin*-A, E and K, and the other produces five fragments including *Hin*-G and J (Lee, Danna, Smith & Nathans, unpublished observations). From the position of these fragments in the cleavage map shown in Figure 3 and further analysis of digest products, we can deduce that the first enzyme fraction (designated *Hind* III) cleaves SV40 DNA at *Hin* sites 4, 6, 7, 9, 10 and 11, whereas the second enzyme fraction (designated *Hind* II) cleaves at *Hin* sites 1, 2, 3, 5 and 8.

The availability of ordered sets of specific fragments of SV40 DNA is proving useful in analyzing functions of the genome. For example, the initiation site for DNA replication has been located within fragment *Hin*-C close to site *Hin*-7 (at about 0·67 of a map unit) and the termination of replication has been mapped within fragment *Hin*-G (at about 0·15 of a map unit), thus allowing the conclusion that SV40 DNA replication is bidirectional (Danna & Nathans, 1972) and proceeds at about equal rates in the two replication arms. (These values are in close agreement with those reported by Fareed *et al.* (1972) who used electron micrograph length measurements to localize the origin of replication.) *Hin* fragments have also been used to localize the genes in the SV40 chromosome that are transcribed early and late in infection and to determine the direction of transcription (Khoury *et al.*, 1973). Similarly, deletions in the SV40 chromosome (Yoshiike, 1968) and substitutions with cellular DNA (Lavi & Winocour, 1972; Tai *et al.*, 1972), that occur during serial passage of SV40 at high multiplicity, are being mapped in relation to the position of restriction endonuclease cleavage sites (Brockman *et al.*, 1973). Finally, the SV40 cleavage map provides a framework for relating the results of nucleotide sequence analyses of individual fragments to the over-all structure of the genome.

We thank Dr Theresa Lee for her help in some of the experiments reported.

This research was supported by grants from the National Cancer Institute, United States Public Health Service, and the Whitehall Foundation, Inc. One of us (G. H. S.) was supported by training grant no. GM00624 from the United States Public Health Service and another author (K. J. D.) was a predoctoral fellow of the United States Public Health Service.

The results reported in this paper were presented at the Cold Spring Harbor Symposium on Tumor Viruses in August, 1972.

REFERENCES

Adler, S. P. & Nathans, D. (1970). *Fed. Proc.* **29**, 725.
Brockman, W. W., Lee, T. N. H. & Nathans, D. (1973). *Virology*, in the press.
Danna, K. J. & Nathans, D. (1971). *Proc. Nat. Acad. Sci., U.S.A.* **68**, 2913.
Danna, K. J. & Nathans, D. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3097.
DeWachter, R. & Fiers, W. (1971). In *Methods in Enzymology* (Colowick, S. P. & Kaplan, N. O., eds), vol. 21, pp. 167–178, Academic Press, New York.
Fairbanks, G., Jr, Levinthal, C. & Reeder, R. H. (1965). *Biochem. Biophys. Res. Commun.* **20**, 393.
Fareed, G. C., Garon, C. F. & Salzman, N. P. (1972). *J. Virol.* **10**, 484.
Gromkova, R. & Goodgal, S. H. (1972). *J. Bact.* **109**, 987.
Hirt, B. (1967). *J. Mol. Biol.* **26**, 365.
Kelly, T. J., Jr & Smith, H. O. (1970). *J. Mol. Biol.* **51**, 393.
Khoury, G., Martin, M. A., Lee, T. N. H., Danna, K. J. & Nathans, D. (1973). *J. Mol. Biol.* **78**, 377.
Lavi, S. & Winocour, E. (1972). *J. Virol.* **9**, 309.
Maizel, J. V., Jr. (1971). In *Methods in Virology* (Maramorosch, K. & Koprowski, H., eds), vol. 5, pp. 180–247, Academic Press, New York.
Morrow, J. & Berg, P. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3365.
Mulder, C. & Delius, H. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3215.
Nathans, D. & Danna, K. J. (1972). *J. Mol. Biol.* **64**, 515.
Reid, M. S. & Bieleski, R. L. (1968). *Anal. Biochem.* **22**, 374.
Sack, G. H., Jr. & Nathans, D. (1973). *Virology*, **51**, 517.
Smith, H. O. & Wilcox, K. (1970). *J. Mol. Biol.* **51**, 379.
Tai, H. T., Smith, C. A., Sharp, P. A. & Vinograd, J. (1972). *J. Virol.* **9**, 317.
Yoshiike, K. (1968). *Virology*, **34**, 391.
Yoshimoro, R. N. (1971). Ph.D. Thesis, The University of California Medical Center, San Francisco, U.S.A.